

Relational Databases and Bell's Theorem

Samson Abramsky

September 3, 2012

Abstract

Our aim in this paper is to point out a surprising formal connection, between two topics which seem on face value to have nothing to do with each other: relational database theory, and the study of non-locality and contextuality in the foundations of quantum mechanics. We shall show that there is a remarkably direct correspondence between central results such as Bell's theorem in the foundations of quantum mechanics, and questions which arise naturally and have been well-studied in relational database theory.

1 Introduction

Our aim in this paper is to point out a surprising formal connection, between two topics which seem on face value to have nothing to do with each other:

- Relational database theory.
- The study of non-locality and contextuality in the foundations of quantum mechanics.

We shall show, using the unified treatment of the latter developed in [1], that there is a remarkably direct correspondence between central results such as Bell's theorem in the foundations of quantum mechanics, and questions which arise naturally and have been well-studied in relational database theory.

The present paper is meant to be an introduction to these two topics, emphasizing their common content, presented in a manner which hopefully will be accessible to readers without prior knowledge of either.

How should this unexpected connection be interpreted? One idea is that the notion of contextuality is rather fundamental, and we can see some outlines of a common 'logic of contextuality' arising from this appearance of common structure in very different settings.

Ideally, some deeper connections can also be found, leading to interesting transfers of results and methods. The present paper may at least serve to point the way in this direction.

2 Relational Databases

2.1 Review of basic notions

We shall begin by reviewing some basic notions of relational database theory.

We start with an example to show the concrete scenario which is to be formalized.

Example

Consider the following data table:

branch-name	account-no	customer-name	balance
Cambridge	10991-06284	Newton	£2,567.53
Hanover	10992-35671	Leibniz	€11,245.75
...

Let us anatomize this table. There are a set of *attributes*,

$$\{\mathbf{branch-name}, \mathbf{account-no}, \mathbf{customer-name}, \mathbf{balance}\}$$

which name the columns of the table. The entries in the table are ‘tuples’ which specify a value for each of the attributes. The table is a set of such tuples. A database will in general have a set of such tables, each with a given set of attributes. The *schema* of the database — a static, syntactic specification of the kind of information which can reside in the database — is given by specifying the set of attributes for each of the tables. The state of the database at a given time will be given by a set of tuples of the appropriate type for each of the tables in the schema.

We now proceed to formalize these notions.

We fix some set \mathcal{A} which will serve as a universe of attributes. A database schema Σ over \mathcal{A} is a family $\Sigma = \{A_1, \dots, A_k\}$ of finite subsets of \mathcal{A} .

At this — surprisingly early! — point, we come to an interesting juncture. The standard approach in much of the current relational database literature is to formalize the notion of tuple as an ordered n -tuple in D^n for some set D of data values; a relation is then a subset of D^n . This is motivated by the desire to make the connection to the standard notion of relational structure in first-order logic as direct as possible. This choice creates a certain distance between the formal notion of relation, and the informal notion of table; in practice this is not a problem.

For our purposes, however, we wish to make a different choice: we shall formalize the notion of tuple, and hence of relation, in a fashion which directly reflects the informal notion. As we shall see, this will have both mathematical and conceptual advantages for our purposes. At the same time, there will be no real problem in relating this formalism to the standard one in the literature. Note that the style of formalization we shall use is also commonly found in the older literature on relational databases, see e.g. [25].

We shall assume that for each $a \in \mathcal{A}$ there is a set D_a of possible data values for that attribute. Thus for example the possible values for **customer-name** should be character strings, perhaps with some lexical constraints; while for **balance** the values should be pairs (**currency**, **amount**), where **currency** comes from some fixed list (£, €, ...), and **amount** is a number. These correspond to *domain integrity constraints* in the usual database terminology.

Given $A \in \Sigma$, we define the set of A -tuples to be $\prod_{a \in A} D_a$. Thus an A -tuple is a function which assigns a data value in D_a to each $a \in A$.

In our example above, the first tuple in the table corresponds to the function

$$\begin{aligned} \{\mathbf{branch-name} \mapsto \text{Cambridge}, \mathbf{account-no} \mapsto 10991-06284, \mathbf{customer-name} \mapsto \text{Newton}, \\ \mathbf{balance} \mapsto \text{£2,567.53}\} \end{aligned}$$

A relation of type A is a set of A -tuples. Given a schema Σ , an *instance* of the schema, representing a possible state of the database, is given by specifying a relation of type A for each $A \in \Sigma$.

Operations on relations

We consider some of the fundamental operations on relations, which play a central rôle in relational databases. Firstly, relations of type A live in the powerset $\mathcal{P}(\prod_{a \in A} D_a)$, which is a boolean algebra; so boolean operations such as union, intersection, and set difference can be applied to them.

At this point, we note that our assumption that each attribute $a \in \mathcal{A}$ has a given set of possible data values is not entirely innocent. It means that boolean operations such as set complement

of relations have an ‘absolute’, *i.e.* instance-independent meaning. By contrast, in the standard view there is no fixed universe of values for instances. An instance specifies its own universe U as well as the interpretations of the database relations as sets of ordered tuples $R_i \subseteq U^{n_i}$ over this universe.

While this difference in formalization could have some interesting consequences, we shall not pursue this here.

Next, we consider the operation of projection. In the language of A -tuples, projection is *function restriction*. That is, given an A -relation R , and a subset $B \subseteq A$, we define:

$$R|B := \{t|B : t \in R\}.$$

Here, since $t \in \prod_{a \in A} D_a$, $t|B$ just means restriction of the function t to B , which is a subset of its domain. This operation is then lifted pointwise to relations.

Now we consider the independent combination of relations, which is cartesian product in the standard formalism. The representation of tuples as functions leads to a ‘logarithmic shift’ in the representation¹, whereby this operation is represented by *disjoint union* of attribute sets.

Given an A -relation R and a B -relation S , we form the disjoint union $A \sqcup B$, and the $A \sqcup B$ -relation

$$R \otimes S := \{t \in \prod_{a \in A \sqcup B} D_a : t|A \in R \wedge t|B \in S\}.$$

Of course, as concrete sets A and B may overlap. We can force them to be disjoint by ‘tagging’ them appropriately, e.g.

$$A \sqcup B := \{0\} \times A \cup \{1\} \times B.$$

The minor housekeeping details of such tagging can safely be ignored.² We shall henceforth do so without further comment.

This is only a subset of the operations available in standard relational algebra [25]. A more complete discussion could be given in the present setting, but this will suffice for our purposes.

2.2 The functorial view

We shall now show how the relational database formalism, in the style we have developed it, has a direct expression in functorial terms. This immediately brings a great deal of mathematical structure into play, and will allow us to relate some important database notions to concepts of much more general standing.

We shall assume the rudiments of the language of categories, functors and natural transformations. All the background we shall need is covered in the charming (and succinct) text [24].

We shall consider the partial order **Att** of finite subsets of \mathcal{A} , ordered by inclusion, as a category.

We shall define a functor $\mathcal{T} : \mathbf{Att}^{\text{op}} \longrightarrow \mathbf{Set}$ where $\mathcal{T}(A)$ is the set of A -tuples. Formally, we define

$$\mathcal{T}(A) := \prod_{a \in A} D_a,$$

and if $A \subseteq B$, we define the *restriction map* $\rho_A^B : \mathcal{T}(B) \longrightarrow \mathcal{T}(A)$ by

$$\rho_A^B : t \mapsto t|A.$$

It is easy to verify functoriality of \mathcal{T} , which means that, whenever $A \subseteq B \subseteq C$,

$$\rho_A^B \circ \rho_B^C = \rho_A^C,$$

and also that $\rho_A^A = \text{id}_A$. Thus \mathcal{T} is a *presheaf*, and restriction is exactly function restriction.

¹Think of $2^a 2^b = 2^{a+b}$, and hence $\log(xy) = \log(x) + \log(y)$.

²The relevant result is the coherence theorem for monoidal categories [19].

We also have the covariant powerset functor $\mathcal{P} : \mathbf{Set} \longrightarrow \mathbf{Set}$, which acts on functions by direct image: if $f : X \longrightarrow Y$, then

$$\mathcal{P}f : \mathcal{P}X \longrightarrow \mathcal{P}Y :: S \mapsto \{f(x) : x \in S\}.$$

We can compose \mathcal{P} with \mathcal{T} to obtain another presheaf

$$\mathcal{R} := \mathcal{P} \circ \mathcal{T} : \mathbf{Att}^{\text{op}} \longrightarrow \mathbf{Set}.$$

This presheaf assigns the set of A -relations to each set of attributes A ; while the restriction map

$$\rho_A^B : \mathcal{R}(B) \longrightarrow \mathcal{R}(A)$$

is exactly the operation of relation restriction, equivalent to the standard notion of projection in relation algebra, which we defined previously:

$$\rho_A^B : R \mapsto R|A.$$

Natural Join

One of the most important operations in relational algebra is *natural join*. Given an A -relation R and a B -relation S , we define an $(A \cup B)$ -relation $R \bowtie S$:

$$R \bowtie S := \{t \in \prod_{a \in A \cup B} D_a : t|A \in R \wedge t|B \in S\}.$$

We shall now show how this operation can be characterized in categorical terms.

Note firstly that since the powerset is naturally ordered by set inclusion, we can consider \mathcal{R} as a functor

$$\mathcal{R} : \mathbf{Att}^{\text{op}} \longrightarrow \mathbf{Pos}$$

where \mathbf{Pos} is the category of posets and monotone maps. \mathbf{Pos} is order-enriched; given monotone maps $f, g : P \rightarrow Q$, we can define the pointwise order:

$$f \leq g \equiv \forall x \in P. f(x) \leq g(x).$$

Now suppose we are given attribute sets A and B . We consider the following diagram arising from the universal property of product in \mathbf{Set} .

$$\begin{array}{ccccc}
 \mathcal{R}(A) & \xleftarrow{\pi_1} & \mathcal{R}(A) \times \mathcal{R}(B) & \xrightarrow{\pi_2} & \mathcal{R}(B) \\
 & \searrow \rho_A^{A \cup B} & \uparrow \langle \rho_A^{A \cup B}, \rho_B^{A \cup B} \rangle & \nearrow \rho_B^{A \cup B} & \\
 & & \mathcal{R}(A \cup B) & &
 \end{array}$$

Proposition 2.1 *The natural join $\bowtie : \mathcal{R}(A) \times \mathcal{R}(B) \rightarrow \mathcal{R}(A \cup B)$ is uniquely characterized as the left adjoint of $\langle \rho_A^{A \cup B}, \rho_B^{A \cup B} \rangle$; that is, as the unique map satisfying*

$$\text{id}_{\mathcal{R}(A \cup B)} \leq \bowtie \circ \langle \rho_A^{A \cup B}, \rho_B^{A \cup B} \rangle, \quad \langle \rho_A^{A \cup B}, \rho_B^{A \cup B} \rangle \circ \bowtie \leq \text{id}_{\mathcal{R}(A) \times \mathcal{R}(B)}.$$

The fact that in general a relation $R \in \mathcal{R}(A \cup B)$ satisfies only

$$R \subseteq R|A \bowtie R|B,$$

with strict inclusion possible, corresponds to the fact that natural join is in general a ‘lossy’ operation. Lossless joins correspond exactly to the case when equality holds.

2.3 The sheaf-theoretic view

We shall now show, building on the presheaf structure described in the previous sub-section, how a number of important database notions can be interpreted geometrically, in the language of sheaves and presheaves.

Schemas as covers and gluing conditions

We shall interpret a schema $\Sigma = \{A_1, \dots, A_k\}$ of finite subsets of \mathcal{A} as a *cover*. That is, we think of the attribute sets A_i as ‘open sets’ expressing some local information in the sense of related clusters of attributes; these sets cover $A := \bigcup_{i=1}^k A_i$, the global set of attributes for the schema. Conversely, we think of the global set A as being decomposed into the local clusters A_i ; which is exactly the standard point of view in databases.

The basic idea of sheaf theory is to analyze the passage from local to global behaviour in mathematical structures. A number of important notions in databases have exactly this character, and can be described naturally in sheaf-theoretic terms.

An instance (R_1, \dots, R_k) of a schema Σ is given by specifying a relation $R_i \in \mathcal{R}(A_i)$ for each $A_i \in \Sigma$. In sheaf-theoretic language, this is a family of local sections, defined over the open sets in the cover. A central issue in geometric terms is whether we can glue these local sections together into a global section defined over $A := \bigcup_{i=1}^k A_i$.

More precisely, we can ask whether there exists a relation $R \in \mathcal{R}(A)$ such that $R|A_i = R_i$, $i = 1, \dots, k$.

This has been studied as an algorithmic question in database theory; it is shown in [14] that it is NP-complete.

Note that a *necessary condition* for this to hold is that, for all $i, j = 1, \dots, k$:

$$R_i|A_i \cap A_j = R_j|A_i \cap A_j. \tag{1}$$

Indeed, if such an R exists, then

$$R_i|A_i \cap A_j = (R|A_i)|A_i \cap A_j = R|A_i \cap A_j,$$

using the functoriality of restriction, and similarly for $R_j|A_i \cap A_j$.

We shall say that a database instance (R_1, \dots, R_k) for which this condition (1) holds has consistent projections, and refer to the family of relations in the instance as a *compatible family*.

These notions can be generalized to apply to any presheaf. If the gluing condition can *always* be satisfied, for any cover and any family of compatible elements, and moreover there is a *unique* element which satisfies it, then the presheaf is a sheaf.

It is of course a well-known fact of life in databases, albeit expressed in a different language, that our relational presheaf \mathcal{R} is *not* a sheaf.

In fact, we have the following:

Proposition 2.2 *An instance (R_1, \dots, R_k) satisfies the gluing condition if and only if there is a universal relation R for the instance.*

Here we take a universal relation for the instance by definition to be a relation defined on the whole set of attributes from which each of the relations in the instance can be recovered by projection. This notion, and various related ideas, played an important rôle in early developments in relational database theory; see e.g. [21, 9, 18, 20, 25].

Thus the standard notion of universal relation in databases corresponds exactly to the standard notion of solution to the gluing condition in sheaf theory, for the particular case of the relational presheaf \mathcal{R} .

It is also standard that a universal relation need not exist in general, and even if it exists, it need not be unique. There is a substantial literature devoted to the issue of finding conditions under which these properties do hold.

There is a simple connection between universal relations and lossless joins.

Proposition 2.3 *Let (R_1, \dots, R_k) be an instance for the schema $\Sigma = \{A_1, \dots, A_k\}$. Define $R := \bowtie_{i=1}^k R_i$. Then a universal relation for the instance exists if and only if $R|_{A_i} = R_i$, $i = 1, \dots, k$, and in this case R is the largest relation in $\mathcal{R}(\bigcup_i A_i)$ satisfying the gluing condition.*

Proof We note that, if a relation S satisfies $S|_{A_i} = R_i$, $i = 1, \dots, k$, then $S \subseteq \bowtie_{i=1}^k R_i$ by the adjoint property of the natural join. Moreover, since projection is monotone, in this case $R_i \subseteq S|_{A_i} \subseteq (\bowtie_{i=1}^k R_i)|_{A_i} \subseteq R_i$. \square

There are further categorical aspects of relational databases which it might prove interesting to pursue. In particular, one can define categories of schemas and of instances and their morphisms, and the construction of colimits in these categories may be applicable to issues of data integration. However, we shall not pursue these ideas here. Instead, we will turn to a natural generalization of relational databases which arises rather effortlessly from the formalism we have developed to this point.

3 Algebraic Databases

We begin by revisiting the definition of the relational presheaf \mathcal{R} in terms of the covariant powerset functor \mathcal{P} . An alternative presentation of subsets is in terms of *characteristic functions*. That is, we have the familiar isomorphism $\mathcal{P}(X) \cong \mathbf{2}^X$, where $\mathbf{2} := \{0, 1\}$ is the 2-element boolean algebra.

We can also use this representation to define the functorial action of powerset. Given $s : X \rightarrow \mathbf{2}$ and $f : X \rightarrow Y$, we define $f^*(s) : Y \rightarrow \mathbf{2}$ by

$$f^*(s) : y \mapsto \bigvee_{f(x)=y} s(x). \quad (2)$$

It is easy to see that this is equivalent to

$$f^*(s)(y) = 1 \iff \exists x \in S. f(x) = y.$$

Here S is the subset of X whose characteristic function is s .

The advantage of this ‘matrix’ style of definition of the powerset is that it can immediately be generalized rather widely. There is a minor caveat. In the above definition, we used the fact that $\mathbf{2}$ is a *complete* boolean algebra, since there was no restriction on the cardinality of the preimages of f . In the database context, of course, all sets are typically finite.³ We shall enforce a finiteness condition explicitly in our general definition.

We recall that a *commutative semiring* is a structure $(R, +, 0, \cdot, 1)$, where $(R, +, 0)$ and $(R, \cdot, 1)$ are commutative monoids, and moreover multiplication distributes over addition:

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Many examples of commutative semirings arise naturally in Computer Science: we list a few of the most common.

- The reals

$$(\mathbb{R}, +, 0, \times, 1).$$

More generally, any commutative ring is a commutative semiring.

³The sets of data values D_a may be infinite, but only finitely many values will appear in a database instance.

- The non-negative reals

$$(\mathbb{R}_{\geq 0}, +, 0, \times, 1).$$

- The booleans

$$\mathbf{2} = (\{0, 1\}, \vee, 0, \wedge, 1).$$

More generally, *idempotent* commutative semirings are exactly the distributive lattices.

- The min-plus semiring

$$(\mathbb{R}_{\geq 0} \cup \{\infty\}, \min, \infty, +, 0).$$

We also note the rôle played by *provenance semirings* in database theory [11, 6, 8].

We fix a semiring R . Given a set X , the *support* of a function $v : X \rightarrow R$ is the set of $x \in X$ such that $v(x) \neq 0$. We write $\text{supp}(v)$ for the support of v . We shall write $\mathcal{V}_R(X)$ for the set of functions $v : X \rightarrow R$ of *finite* support. We shall write $\mathcal{D}_R(X)$ for the subset of $\mathcal{V}_R(X)$ of those functions $d : X \rightarrow R$ such that

$$\sum_{x \in X} d(x) = 1.$$

Note that the finite support condition ensures that this sum is well-defined.

We shall refer to elements of $\mathcal{V}_R(X)$ as R -valuations on X , and of $\mathcal{D}_R(X)$ as R -distributions.

We consider a few examples:

- If we take $R = \mathbf{2}$, then $\mathcal{V}_R(X)$ is the set of finite subsets of X , and $\mathcal{D}_R(X)$ is the set of finite non-empty subsets.
- If we take $R = (\mathbb{R}_{\geq 0}, +, 0, \times, 1)$, then $\mathcal{D}_R(X)$ is the set of discrete (finite-support) probability distributions on X .

Algebraically, $\mathcal{V}_R(X)$ is the free R -semimodule over the set X [10].

These constructions extend to functors on **Set**. Given $f : X \rightarrow Y$, we define

$$\mathcal{V}_R(f) : \mathcal{V}_R(X) \rightarrow \mathcal{V}_R(Y) :: v \mapsto [y \mapsto \sum_{f(x)=y} v(x)].$$

This restricts to \mathcal{D}_R in a well-defined fashion. Taking $R = \mathbf{2}$, we see that $\mathcal{V}_R(f)$ is exactly the direct image of f , defined as in (2).

In fact, \mathcal{V}_R and \mathcal{D}_R are both monads on **Set**; \mathcal{V}_R is commutative, and \mathcal{D}_R is affine [17, 16].

We can now generalize databases from the standard relational case to ‘relations valued in a semiring’ by replacing \mathcal{P} by \mathcal{V}_R (or \mathcal{D}_R) in our definition of \mathcal{R} ; that is, we take $\mathcal{R} := F \circ \mathcal{T}$, where F is \mathcal{V}_R or \mathcal{D}_R for some commutative semiring R . We recover the standard notion exactly when $R = \mathbf{2}$. In the case where $R = (\mathbb{R}_{\geq 0}, +, 0, \times, 1)$ and $F = \mathcal{D}_R$, we obtain a notion of probabilistic database, where each relation specifies a probability distribution over the set of tuples for its attribute-set.

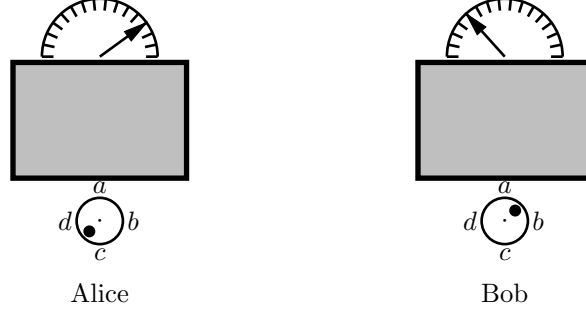
We also note an important connection between probabilistic and relational databases. We can always pass from a probabilistic to a relational instance by taking the *support* of the distribution. Algebraically, this corresponds to mapping all positive probabilities to 1; this is in fact the action of the unique semiring homomorphism from the non-negative reals to the booleans.

In general, many natural properties of databases will be *preserved* by this homomorphic mapping. This means that if we show that such a property is *not* satisfied by the support, we can conclude that it is not satisfied by the probabilistic instance. Thus we can leverage negative results at the relational level, and lift them to the probabilistic setting.

4 From databases to observational scenarios

We shall now offer an alternative interpretation of the relational database formalism, with a very different motivation. This will expose a surprising connection between database theory, and on face value a completely different topic, namely Bell's theorem in the foundations of quantum mechanics [5].

Our starting point is the idealized situation depicted in the following diagram.



There are several agents or experimenters, who can each select one of several different measurements to perform, and observe one of several different outcomes. These agents may or may not be spatially separated. When a system is prepared in a certain fashion and measurements are selected, some corresponding outcomes will be observed. These individual occurrences or ‘runs’ of the system are the basic events. Repeated runs allow relative frequencies to be tabulated, which can be summarized by a probability distribution on events for each selection of measurements. We shall call such a family of probability distributions, one for each choice of measurements, an *empirical model*.

As an example of such a model, consider the following table.

		(0, 0)	(1, 0)	(0, 1)	(1, 1)
a	b	1/2	0	0	1/2
a'	b	3/8	1/8	1/8	3/8
a	b'	3/8	1/8	1/8	3/8
a'	b'	1/8	3/8	3/8	1/8

The intended scenario here is that Alice can choose between measurements a and a' , and Bob can choose b or b' ; we assume that these choices are made independently. Thus the *measurement contexts* are

$$\{a, b\}, \quad \{a', b\}, \quad \{a, b'\}, \quad \{a', b'\},$$

and these index the rows of the table. Each measurement has possible outcomes 0 or 1.

Note that, with a small change of perspective, we can see this in database terms. Take the global set of attributes $\mathcal{A} = \{a, a', b, b'\}$, and consider the schema

$$\Sigma := (\{a, b\}, \{a', b\}, \{a, b'\}, \{a', b'\}).$$

For each $a \in \mathcal{A}$, we take $D_a := \{0, 1\}$.

For each $A \in \Sigma$, we have a ‘table’ in the algebraically generalized sense discussed in the previous section. That is, we have a distribution $d_A \in \mathcal{D}_R \circ \mathcal{T}(A)$, where $R = \mathbb{R}_{\geq 0}$ is the semiring of non-negative reals. Thus d_A is a probability distribution on $\mathcal{T}(A)$, the set of A -tuples.

To make a direct connection with standard relational databases, we can pass to the support of the above table, which yields the following:

		(0,0)	(1,0)	(0,1)	(1,1)
a	b	1	0	0	1
a'	b	1	1	1	1
a	b'	1	1	1	1
a'	b'	1	1	1	1

This corresponds to the instance of the schema Σ where for each $A = \{\alpha, \beta\} \in \Sigma \setminus \{a, b\}$, there is the ‘full’ table of all possible tuples:

α	β
0	0
0	1
1	0
1	1

while for $\{a, b\}$ we have the table with only two tuples:

a	b
0	0
1	1

Thus we have a formal passage between empirical models and relational databases. To go further, we must understand how empirical models such as these can be used to draw striking conclusions about the foundations of physics.

5 Empirical models and hidden variables

Most of our discussion is independent of any particular physical theory. However, it is important to understand how quantum mechanics, as our most highly confirmed theory, gives rise to a class of empirical models of the kind we have been discussing.

To obtain such a model, we must provide the following ingredients:

- A quantum state.
- For each of the ‘measurement settings’, which correspond to attributes in database terms, a physical observable or measurable quantity. Each such observable will have a set of associated possible outcomes, which will correspond to the set of data values associated with that attribute.

The ‘statistical algorithm’ of quantum mechanics will then prescribe a probability for each measurement outcome when the given state is measured with that observable.

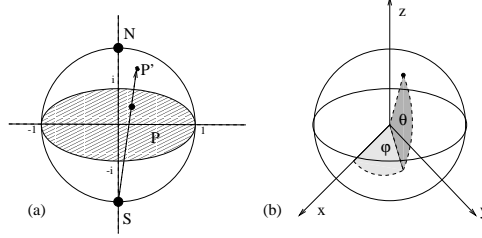
Although we shall not really need the details of this, we briefly recall some basic definitions. For further details, see e.g. [23].

A crash course in qubits

Whereas a classical bit register has possible states 0 or 1, a qubit state is given by a *superposition* of these states. More precisely, a (pure) qubit state is given by a complex linear combination $\alpha_0|0\rangle + \alpha_1|1\rangle$, subject to the normalization constraint $|\alpha_0|^2 + |\alpha_1|^2 = 1$.

Measurement of such a state (in the $|0\rangle, |1\rangle$ basis) is inherently probabilistic; we get $|i\rangle$ with probability $|\alpha_i|^2$.

There is a beautiful geometric picture of this complex 2-dimensional geometry in real three-dimensional space. This is the Bloch sphere representation:



The pure qubit states correspond to points on the surface of the sphere. However, this one-qubit case does not yet provide non-classical resources for information processing. Things get interesting with n -qubit registers

$$\sum_i \alpha_i |i\rangle, \quad i \in \{0, 1\}^n.$$

It is at this point, in particular, that *entanglement phenomena* arise.

A typical example of an entangled state is the Bell state:

$$|00\rangle + |11\rangle$$

We can think of two particles, each with a qubit state, held by Alice and Bob. However, these two particles are entangled. If Alice measures her qubit, then if she gets the answer $|0\rangle$, the state will collapse to $|00\rangle$, and if Bob measures his qubit, he will get the answer $|0\rangle$ with certainty; similarly if the result of Alice's measurement is $|1\rangle$. This non-local effect creates new possibilities for quantum information processing.

Mathematically, compound systems are represented by the *tensor product*, $\mathcal{H}_1 \otimes \mathcal{H}_2$, with typical element

$$\sum_i \lambda_i \cdot \phi_i \otimes \psi_i.$$

Superposition encodes *correlation*.

Entanglement is the physical phenomenon underlying Einstein's 'spooky action at a distance'. Even if the particles are spatially separated, measuring one has an effect on the state of the other.

Bell's achievement was to turn this puzzling feature of quantum mechanics into a theorem: quantum mechanics is *essentially non-local*.

5.1 Bell's theorem

We look again at the empirical model

	(0, 0)	(1, 0)	(0, 1)	(1, 1)
(a, b)	1/2	0	0	1/2
(a, b')	3/8	1/8	1/8	3/8
(a', b)	3/8	1/8	1/8	3/8
(a', b')	1/8	3/8	3/8	1/8

This can be realized in quantum mechanics, using a Bell state

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}},$$

subjected to measurements in the XY -plane of the Bloch sphere, at relative angle $\pi/3$. Systems of this kind have been the subject of extensive experimental investigation, and the predictions of quantum mechanics can be taken to be very highly confirmed.

The question we shall ask, following Bell, is this: Can we explain these empirical findings by a theory which is *local* and *realistic* in the following sense.

- A theory is realistic if it ascribes definite values to all observables for every physical state, independently of the activities of any external observers.
- A theory is local if the outcomes of measurements on spatially separated subsystems depend only on common causal factors. In particular, for space-like separated measurements, the outcomes of the measurements should be independent of each other.

We allow for the fact that there may be salient features in the theory determining the outcomes of measurements of which we are not aware. These features are embodied in the notion of *hidden variable*. Thus we take measurement outcomes to be determined, *given some value of this hidden variable*. Moreover, we assume that this hidden variable acts in a local fashion with respect to spatially separated subsystems.

This gives a general notion of theory which behaves in a fashion broadly consistent with classical physical intuitions. The import of Bell's theorem is exactly that *no such theory can account for the empirical predictions of quantum mechanics*. Hence, given that these predictions are so well-confirmed, we must abandon the classical world-view which underpins the assumptions of local realism.

To give a precise statement of Bell's theorem, we must formalize the notion of local hidden variable theory. We shall give this in a streamlined form, which can be shown to be equivalent to more general definitions which have been considered (see e.g. Theorem 7.1 in [1]).

We shall explain this notion in relation to the Bell table given above. We have a total set of four measurement settings we are considering, two for Alice and two for Bob:

$$\{a, a', b, b'\}.$$

A simultaneous assignment of outcomes (0 or 1) to each of these is given by a function

$$s : \{a, a', b, b'\} \longrightarrow \{0, 1\}.$$

The fact that an (unknown) hidden variable may be affecting the outcome is captured by saying that we have a probability distribution d on the set of all such functions s . Such a probability distribution can be taken to be a canonical form for a hidden variable.

The requirement on this distribution d to be consistent with the empirical data is that, for each of the experimentally accessible combinations of measurement settings

$$\{a, b\}, \quad \{a', b\}, \quad \{a, b'\}, \quad \{a', b'\},$$

the restriction (or marginalization) of d to this set of measurements yields exactly the observed distribution on outcomes from the corresponding row of the table. For example, we must have $d|\{a, b\} = d_1$, where

$$d_1(0, 0) = d_1(1, 1) = 1/2, \quad d_1(0, 1) = d_1(1, 0) = 0.$$

A precise statement of a particular instance of Bell's theorem can now be given as follows:

Proposition 5.1 *There is no distribution d on the whole set of measurements which yields the observable distributions by restriction.*

Proof Assume for a contradiction that such a distribution d exists. It will assign a number $X_i \in [0, 1]$ to each $s_i : \{a, a', b, b'\} \longrightarrow \{0, 1\}$. There are 16 such functions: we enumerate them by viewing them as binary strings, where the j 'th bit indicates the assignment of an outcome to the j 'th measurement, listed as a, a', b, b' .

The requirement that this distribution projects onto the distributions in the empirical model translates into 16 equations, one for each entry in the table. It suffices to consider 4 of these equations:

$$\begin{aligned} X_1 + X_2 + X_3 + X_4 &= 1/2 \\ X_2 + X_4 + X_6 + X_8 &= 1/8 \\ X_3 + X_4 + X_{11} + X_{12} &= 1/8 \\ X_1 + X_5 + X_9 + X_{13} &= 1/8 \end{aligned}$$

Adding the last three equations yields

$$X_1 + X_2 + X_3 + 2X_4 + X_5 + X_6 + X_8 + X_9 + X_{11} + X_{12} + X_{13} = 3/8.$$

Since all these terms must be non-negative, the left-hand side of this equation must be greater than or equal to the left-hand side of the first equation, yielding the required contradiction. \square

This argument seems very specific to the probabilistic nature of the empirical model. However, an important theme in the work on no-go theorems in quantum mechanics is to prove results of this kind in a probability-free fashion [12, 13]. This will bring us directly into the arena of relational databases.

5.2 Hardy's construction

Hardy's construction [13] yields a family of empirical models which can be realized in quantum mechanics in similar fashion to the Bell model. However, these families exhibit a stronger form of non-locality property, which does not depend on the probabilities, but only on the support.

We exhibit an example of a support table arising from Hardy's construction.

	(0, 0)	(1, 0)	(0, 1)	(1, 1)
(a, b)	1	1	1	1
(a', b)	0	1	1	1
(a, b')	0	1	1	1
(a', b')	1	1	1	0

This arises from a probability table by replacing all positive probabilities by 1.

Note that we can view this table as encoding a small relational database, as in our discussion in the previous section. There will be four relation tables in this database, one for each of the above rows. The table corresponding to the first row will have the full set of tuples over $\{0, 1\}$. The tables for the second and third rows will have the form

α	β
0	1
1	0
1	1

while that for the fourth row will have the form

a'	b'
0	0
0	1
1	0

The property which shows the non-locality of this model is the exact relational analogue of the probabilistic version we considered in relation to the Bell model.

Proposition 5.2 *There is no A -relation R , where $A = \{a, a', b, b'\}$, which is consistent with the empirical observable supports; that is, for which $R|\{\alpha, \beta\}$ yields the relational table for all $\{\alpha, \beta\}$, $\alpha \in \{a, a'\}$, $\beta \in \{b, b'\}$.*

In database language, this says exactly that there is no ‘universal relation’ on the whole set of attributes which yields each of the ‘observable relations’ by projection.

Proof We argue similarly to the case of the Bell model, except that we are now working over the Boolean semiring rather than the non-negative reals. The existence of a relation R thus reduces to a Boolean satisfiability problem. An equation $\sum_i X_i = 1$ simply asserts the disjunction of the boolean variables, while an equation $\sum_i X_i = 0$ asserts the conjunction of the negated variables. Again it suffices to consider four of the equations which can be read off the Hardy table:

$$\begin{array}{cccccc}
X_1 & \vee & X_2 & \vee & X_3 & \vee & X_4 \\
\neg X_1 & \wedge & \neg X_3 & \wedge & \neg X_5 & \wedge & \neg X_7 \\
\neg X_1 & \wedge & \neg X_2 & \wedge & \neg X_9 & \wedge & \neg X_{10} \\
\neg X_4 & \wedge & \neg X_8 & \wedge & \neg X_{12} & \wedge & \neg X_{16}
\end{array}$$

Since every disjunct in the first formula appears as a negated conjunct in one of the other three formulas, there is no satisfying assignment. \square

There is a precise sense in which the Hardy result is stronger than the Bell result. In fact, we have the following.

Proposition 5.3 *If an empirical model has a local hidden-variable model in the probabilistic sense, then its support table has a universal relation. Thus failure to have a universal relation implies failure to have a local hidden-variable model in the probabilistic sense.*

Proof This follows simply from the fact that the map from the non-negative reals to the booleans which takes all non-zero elements to 1 is a semiring homomorphism. \square

The converse to this result is false. For example, the support table arising from the Bell model *does* have a universal relation, as can easily be verified.

Note that the Hardy table, and indeed all such tables arising from quantum mechanics, satisfies the compatibility condition which we discussed in Section 2.3. In fact, compatibility corresponds precisely to the physical condition of *no-signalling*, and the fact that quantum models satisfy the condition is exactly the content of the No-Signalling theorem of quantum mechanics. See [1] Section 8 for an extended discussion of this point.

6 No-Go theorems, global sections and universal relations

We shall now develop a more general perspective on the results we have discussed in the previous section.

Following the geometric language we introduced in Section 2.3, we see that the existence of a hidden-variable model is equivalently expressed as the existence of a *global section* which glues together the family of empirical accessible distributions or relations.

Thus non-locality and related no-go results can be understood in terms of *obstructions to the existence of global sections*, a central issue in the pervasive applications of sheaves in geometry, topology, analysis and number theory.

In terms of databases, such results can be understood as expressing obstructions to the existence of universal relations for given instances of the database.

We shall now discuss two further types of no-go results, which can be understood in terms of yet stronger forms of obstruction.

6.1 Strong Contextuality

If we consider the argument for the Hardy construction, it can be understood as saying that there is no relation over the global tuples which ‘covers’ all (and only) the observable tuples. But now suppose we consider a much weaker requirement: we simply ask for *one global tuple* which projects consistently into all the relations in the database instance.

Note that the Hardy model does meet this condition. The global assignment

$$\{a \mapsto 1, a' \mapsto 0, b \mapsto 1, b' \mapsto 0\}$$

does project consistently into the support table for this model. The Bell model similarly meets this condition.

If even this, much weaker requirement *cannot* be met, then we have a much stronger form of no-go theorem. We say that such a situation exhibits *strong contextuality*.

The question now arises: are there models coming from quantum mechanics which are strongly contextual in this sense?

We shall now show that the well-known GHZ models [12], of type $(n, 2, 2)$ for all $n > 2$, are strongly contextual. This will establish a strict hierarchy

$$\text{Bell} < \text{Hardy} < \text{GHZ}$$

of increasing strengths of obstructions to non-contextual behaviour for these salient models.

The GHZ model of type $(n, 2, 2)$ can be specified as follows. We label the two measurements at each part as $X^{(i)}$ and $Y^{(i)}$, and the outcomes as 0 and 1. For each context C , every s in the support of the model satisfies the following conditions:

- If the number of Y measurements in C is a multiple of 4, the number of 1’s in the outcomes specified by s is even.
- If the number of Y measurements is $4k + 2$, the number of 1’s in the outcomes is odd.

A model with these properties can be realized in quantum mechanics, using the GHZ state

$$\frac{|0 \cdots 0\rangle + |1 \cdots 1\rangle}{\sqrt{2}}.$$

Proposition 6.1 *The GHZ models are strongly contextual, for all $n \geq 3$.*

Proof We consider the case where $n = 4k$, $k \geq 1$. Assume for a contradiction that we have a global section $s \in S_e$ for the GHZ model e .

If we take Y measurements at every part, the number of 1 outcomes under the assignment is even. Replacing any two Y ’s by X ’s changes the residue class mod 4 of the number of Y ’s, and hence must result in the opposite parity for the number of 1 outcomes under the assignment. Thus for any $Y^{(i)}$, $Y^{(j)}$ assigned the *same* value, if we substitute X ’s in those positions they must receive *different* values under s . Similarly, for any $Y^{(i)}$, $Y^{(j)}$ assigned different values, the corresponding $X^{(i)}$, $X^{(j)}$ must receive the same value.

Suppose firstly that not all $Y^{(i)}$ are assigned the same value by s . Then for some i, j, k , $Y^{(i)}$ is assigned the same value as $Y^{(j)}$, and $Y^{(j)}$ is assigned a different value to $Y^{(k)}$. Thus $Y^{(i)}$ is also assigned a different value to $Y^{(k)}$. Then $X^{(i)}$ is assigned the same value as $X^{(k)}$, and $X^{(j)}$ is

assigned the same value as $X^{(k)}$. By transitivity, $X^{(i)}$ is assigned the same value as $X^{(j)}$, yielding a contradiction.

The remaining cases are where all Y 's receive the same value. Then any pair of X 's must receive different values. But taking any 3 X 's, this yields a contradiction, since there are only two values, so some pair must receive the same value.

The case when $n = 4k + 2$, $k \geq 1$, is proved in the same fashion, interchanging the parities. When $n \geq 5$ is odd, we start with a context containing one X , and again proceed similarly.

The most familiar case, for $n = 3$, does not admit this argument, which relies on having at least 4 Y 's in the initial configuration. However, for this case one can easily adapt the well-known argument of Mermin using 'instruction sets' [22] to prove strong contextuality. This uses a case analysis to show that there are 8 possible global sections satisfying the parity constraint on the 3 measurement combinations with 2 Y 's and 1 X ; and all of these violate the constraint for the XXX measurement. \square

6.2 The Kochen-Specker theorem

Kochen-Specker-type theorems [15] can be understood as *generic strong contextuality results*. In database terms, they say that, if the database schema has a certain combinatorial structure, then *every instance* satisfying some conditions is strongly contextual. This can be interpreted in the quantum context in such a way that the conditions will be satisfied by *every* quantum state, and hence we obtain a state-independent form of strong contextuality result.

The condition which is typically imposed on the instances, assuming that the possible data values for each attribute lie in $\{0, 1\}$, is that *every tuple contains exactly one 1*. If we think in terms of satisfiability, this corresponds to a 'POSITIVE ONE-IN- k -SAT' condition.

To show that the Kochen-Specker result holds is exactly to show that there is no satisfying assignment for the corresponding set of clauses.

The simplest example of this situation is the 'triangle', *i.e.* the schema with elements

$$\{a, b\}, \{b, c\}, \{a, c\}.$$

However, this example cannot be realized in quantum mechanics [1].

An example which can be realized in quantum mechanics, where \mathcal{A} has 18 elements, and there are 9 sets in the database schema, each with four elements, such that each element of \mathcal{A} is in two of these, appears in the 18-vector proof of the Kochen-Specker Theorem in [7].

U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9
A	A	H	H	B	I	P	P	Q
B	E	I	K	E	K	Q	R	R
C	F	C	G	M	N	D	F	M
D	G	J	L	N	O	J	L	O

Here the schema is $\Sigma = \{U_1, \dots, U_9\}$.

We shall give a simple combinatorial condition on the schema Σ which is implied by the existence of a global section s satisfying the 'POSITIVE ONE-IN- k -SAT' condition. Violation of this condition therefore suffices to prove that no such global section exists.

For each $a \in \mathcal{A}$, we define

$$\Sigma(a) := \{A \in \Sigma : a \in A\}.$$

Proposition 6.2 *If a global section satisfying the condition exists, then every common divisor of $\{|\Sigma(a)| : a \in \mathcal{A}\}$ must divide $|\Sigma|$.*

Proof Suppose there is a global section $s : \mathcal{A} \rightarrow \{0, 1\}$ satisfying the condition. Consider the set $X \subseteq \mathcal{A}$ of those a such that $s(a) = 1$. Exactly one element of X must occur in every $A \in \Sigma$. Hence there is a partition of Σ into the subsets $\Sigma(a)$ indexed by the elements of X . Thus

$$|\Sigma| = \sum_{a \in X} |\Sigma(a)|.$$

It follows that, if there is a common divisor of the numbers $|\Sigma(a)|$, it must divide $|\Sigma|$. \square

For example, if every $a \in \mathcal{A}$ appears in an even number of elements of Σ , while Σ has an odd number of elements, then there is no global section. This corresponds to the ‘parity proofs’ which are often used in verifying Kochen-Specker-type results [7, 27]. For example, in the 18-attribute schema with 9 relations given above, each attribute appears in two relations in the schema; hence the argument applies.

For further discussion of these ideas, including connections with graph theory, see [1].

7 Further Directions

We mention some further directions for developing the connections between databases and the study of non-locality and contextuality in quantum mechanics.

- We may consider conditions on the database schema which guarantees that global sections can be found. The important notion of *acyclicity* in database theory [4] seems relevant here. We also note on the probabilistic side a result by Vorob’ev [26] (motivated by game theory), which in effect gives necessary and sufficient combinatorial conditions on a schema for *any* assignment of probability distributions on the tuples for each relation in the schema to have a global section; that is, for a universal relation in the probabilistic sense to always exist for any probabilistic instance of the database.
- A logical approach to Bell inequalities in terms of logical consistency conditions is developed in [2]. It would be interesting to interpret and apply this notion of Bell inequalities in the database context.
- The tools of sheaf cohomology are used to characterize the obstructions to global sections in a large family of cases in [3]. In principle, these sophisticated tools can be applied to databases. There may be interesting connections with acyclicity in the database sense.

Acknowledgements Discussions with and detailed comments by Phokion Kolaitis are gratefully acknowledged. Leonid Libkin also gave useful feedback. This paper was written while in attendance at the program on ‘Semantics and Syntax: the legacy of Alan Turing’ at the Isaac Newton Institute, Cambridge, April–May 2012.

References

- [1] S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(2011):113036, 2011.
- [2] S. Abramsky and L. Hardy. Logical Bell Inequalities. *Arxiv preprint arXiv:1203.1352*, 2012.
- [3] S. Abramsky, S. Mansfield, and R.S. Barbosa. The cohomology of non-locality and contextuality. In *Proceedings of QPL 2011*. EPTCS, 2012. Available as arXiv:1111.3620v1.
- [4] C. Beeri, R. Fagin, D. Maier, and M. Yannakakis. On the desirability of acyclic database schemes. *Journal of the ACM (JACM)*, 30(3):479–513, 1983.
- [5] J.S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3):195–200, 1964.

- [6] P. Buneman and W.C. Tan. Provenance in databases. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, pages 1171–1173. ACM, 2007.
- [7] A. Cabello, J.M. Estebaranz, and G. García-Alcaine. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4):183–187, 1996.
- [8] James Cheney, Laura Chiticariu, and Wang Chiew Tan. Provenance in databases: Why, how, and where. *Foundations and Trends in Databases*, 1(4):379–474, 2009.
- [9] R. Fagin, A.O. Mendelzon, and J.D. Ullman. A simplified universal relation assumption and its properties. *ACM Transactions on Database Systems (TODS)*, 7(3):343–360, 1982.
- [10] J.S. Golan. *Semirings and their Applications*. Springer, 1999.
- [11] T.J. Green, G. Karvounarakis, and V. Tannen. Provenance semirings. In *Proceedings of the twenty-sixth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 31–40. ACM, 2007.
- [12] D.M. Greenberger, M.A. Horne, and A. Zeilinger. Going beyond Bell’s theorem. In M. Kafatos, editor, *Bell’s Theorem, Quantum Theory, and Conceptions of the Universe*, pages 69–72. Kluwer, 1989.
- [13] L. Hardy. Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories. *Physical Review Letters*, 68(20):2981–2984, 1992.
- [14] P. Honeyman, R.E. Ladner, and M. Yannakakis. Testing the universal instance assumption. *Information Processing Letters*, 10(1):14–19, 1980.
- [15] S. Kochen and E.P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967.
- [16] A. Kock. Monads on symmetric monoidal closed categories. *Archiv der Mathematik*, 21(1):1–10, 1970.
- [17] A. Kock. Bilinearity and cartesian closed monads. *Math. Scand*, 29:161–174, 1971.
- [18] H.F. Korth, G.M. Kuper, J. Feigenbaum, A. Van Gelder, and J.D. Ullman. SYSTEM/U: a database system based on the universal relation assumption. *ACM Transactions on Database Systems (TODS)*, 9(3):331–347, 1984.
- [19] S. Mac Lane. *Categories for the working mathematician*, volume 5. Springer Verlag, 1998.
- [20] D. Maier and J.D. Ullman. Maximal objects and the semantics of universal relation databases. *ACM Transactions on Database Systems (TODS)*, 8(1):1–14, 1983.
- [21] D. Maier, J.D. Ullman, and M.Y. Vardi. On the foundations of the universal relation model. *ACM Transactions on Database Systems (TODS)*, 9(2):283–308, 1984.
- [22] N.D. Mermin. Quantum mysteries revisited. *Am. J. Phys*, 58(8):731–734, 1990.
- [23] M.Q.C. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [24] B.C. Pierce. *Basic category theory for computer scientists*. The MIT Press, 1991.
- [25] J.D. Ullman. *Principles of database systems*. Prentice Hall, 1983.
- [26] N.N. Vorob’ev. Consistent families of measures and their extensions. *Theory of Probability and its Applications*, 7:147, 1962.
- [27] M. Waegell and P.K. Aravind. Parity proofs of the Kochen-Specker theorem based on the 24 rays of Peres. *Arxiv preprint arXiv:1103.6058v1*, 2011.